

Theoretical Design Methodology for Practical Interconnection Networks

Ryota Yasudo

A dissertation submitted in partial fulfillment of
the requirements for the degree of

DOCTOR OF PHILOSOPHY

School of Science for Open and Environmental Systems
Graduate School of Science and Technology
Keio University

January 2019

報告番号	Ⓐ 乙 第	号	氏 名	安戸 僚汰
主 論 文 題 名 : Theoretical Design Methodology for Practical Interconnection Networks (実用的相互結合網のための理論的設計方法論)				
<p>(内容の要旨)</p> <p>高性能計算のプラットフォームやハイエンド・データセンター向けの並列アプリケーションにおいては、端点間の通信遅延が重大な問題である。一方ではコンピュータ・アーキテクトらが相互結合網を実験的あるいは経験的に設計しようと試みてきた。他方では理論研究者がコンピュータのネットワークをモデル化し、その性質を理論的に研究してきたが、そのモデルは正確に実際のシステムの性質を捉えているとはいえない。そこで本論文では、理論的な研究と実践的な研究の間に存在するギャップを埋めるために高性能な相互結合網を設計するための新しい手法を確立する。特に本論文はグラフ理論、複雑ネットワーク科学、ブロックデザインの理論などの知識を従来の相互結合網の工学的研究と共に活用する。</p> <p>まずホストスイッチグラフと呼ばれる新しいグラフを導入する。このグラフは最大次数 1 のホスト頂点と最大次数 r のスイッチ頂点を持ち、実用的な並列分散計算機システムのトポロジーを表現する。次いで、高性能な相互結合網を設計する上で重要となる指標であるホスト間平均経路長と二分幅について議論する。特に、r ポートスイッチを使って特定の数のホストをつなぐホストスイッチグラフを、ホスト間平均経路長が小さくて二分幅が大きくなるように構成する方法を探究する。この探究によって、ホスト間平均経路長が最小となるスイッチ数は数学的に予測ができ、二分幅が最大になるようなスイッチ数は実験的に予測できることが明らかとなる。それらの予測に基づいて、ホストスイッチグラフを探索する乱択アルゴリズムを提案する。その後得られたグラフを相互結合網のトポロジーに適用し、典型的なトポロジーと比較を行う。トーラス、ドラゴンフライ、ファットツリーと比較した結果、提案するネットワークはより小さい電力・コストでより高い性能を得ることを示す。</p> <p>さらに、スイッチのポート数を増やすだけでなくホストにポートを付け加えることによって通信遅延を削減する手法を提案する。そのために、ホストスイッチグラフをマルチポート・ホストが表現できるように拡張する。従来マルチポート・ホストはリンク集約 (LA) , ネットワークの複製 (ND) に使われてきたが、これらはホップ数を減らす</p>				

主 論 文 要 旨

No.

ことはない。そこで本論文ではホップ数を削減するためにホスト-スイッチマッピングの置換を提案する。この手法は従来の LA と ND に適用可能であり、それぞれに適用したものを p-LA, p-ND と呼ぶ。それに加えて、ブロックデザイン的一种である有限射影平面の応用を提案し, PP と呼ぶ。提案手法は任意のトポロジーに適用可能のため, 既存トポロジーを直接用いることができる。実験として五つの設計 (LA, ND, p-LA, p-ND, PP) を乱択アルゴリズムによる最適トポロジー, トーラス, ドラゴンフライ, ファットツリーに適用し, 設計複雑性, ホップ数, 二分幅, コスト, ルーティングテーブルのサイズ, メッセージパッシングインタフェース (MPI) によるアプリケーション性能を評価する。結果として, 提案手法 (p-LA, p-ND, PP) が二分バンド幅を増やしコストとホップ数を減らすことを示す。特に PP はコストの面で効率が良く, ホップ数を効果的に減らすことができ, その効果は乱択アルゴリズムによる最適トポロジーとファットツリーで顕著である。

Thesis Abstract

No. _____

Registration Number	<input checked="" type="checkbox"/> "KOU" <input type="checkbox"/> "OTSU" No. _____ *Office use only	Name	Ryota Yasudo
Thesis Title Theoretical Design Methodology for Practical Interconnection Networks			
Thesis Summary <p>End-to-end network latency is a concern for parallel applications in high-performance computing platforms and high-end data centers. On one hand, computer architects have tried to design interconnection networks experimentally and empirically. On the other hand, theoretists have tried to model computer networks and then studied their properties theoretically. However, the model does not sufficiently capture real computer systems. This dissertation aims at establishing a novel method for designing high-performance network topologies to bridge a gap between the theoretical and practical studies. In particular, we make use of the knowledge of graph theory, network science, and design theory, as well as the research on interconnection networks.</p> <p>We firstly present a novel graph called a host-switch graph, which consists of host vertices and switch vertices with maximum degree 1 and r, respectively. This graph represents a network topology of a practical parallel/distributed computer system with host computers connected by r-port switches. We then discuss important metrics for designing high-performance interconnection networks: the host-to-host average shortest path length (h-ASPL) and the bisection width (BiW). In particular, we explore a method for constructing host-switch graphs with low h-ASPL and high BiW that connect the fixed number of hosts via any number of r-port switches. We demonstrate that the number of switches that provides the minimum h-ASPL can mathematically be approximated, and the minimum number of switches that provides a certain BiW can experimentally be approximated. On the basis of the approximations, we propose a randomized algorithm for searching host-switch graphs. We then apply the graphs to interconnection networks and compare them with typical network topologies. As compared with the torus, Dragonfly, and Fat-tree, our networks attain higher performance and smaller power and costs.</p> <p>Furthermore, we propose adding ports to a host as a method for reducing the network latency as well as increasing the number of ports of a switch. To this end, we extend a host-switch graph so that it represents multi-port hosts. Multi-port hosts are conventionally used for link aggregation (LA) and network duplication (ND), but they do not reduce the hop count. We hence propose the permutation of host-switch mapping for reducing the hop count. It can be applied to LA and ND, and we label the obtained networks p-LA and p-ND, respectively. In addition, we propose the application of a finite projective plane (an instance of block designs) and label it PP. Our methods can be applied to arbitrary topologies, and thus we can directly use any existing topologies. We evaluate five designs above (LA, N</p>			

Thesis Abstract

No. _____

D, p-LA, p-ND, and PP) for randomly-optimized, torus, Dragonfly, and Fat-tree topologies in terms of the design complexity, the hop count, the bisection width, costs, the size of routing tables, and simulated message passing interface (MPI) performance. Our results demonstrate that our methods (p-LA, p-ND, and PP) reduce the hop count while increasing the bisection width and costs for every topology. In particular, we demonstrate that PP is a cost-effective method for reducing the hop count, especially for randomly optimized and Fat-tree topologies.